Please replace the paragraph beginning at page 3, line 15 with the following paragraph:

The use of relationships among words can be exploited for image retrieval as taught by Y. Alp Aslandogan, C. Thier, C. T. Yu, J. Zou, N. Rishe in the paper entitled "Using Semantic Contents and WordNet in Image Retrieval," published in *Proc. of the 20th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 286-295, 1997. The system allows the similarity searching of images based on the semantic entity-relationship-attribute descriptions of the image content. WORDNET is used for expanding the query or database for matching. WORDNET is a registered trademark of Trustees of Princeton University, Princeton, New Jersey. The WORDNET system, taught by G. A. Miller in an article entitled "WordNet: A Lexical Database for English," published in *Communication of the ACM*, Vol. 38, No. 11, pp. 39-41, Nov. 1995, incorporated herein by reference, is a graphical network of concepts and associated words in which the relationships among concepts are governed by the form and meaning of the words. However, WORDNET and other textual representations of knowledge do not sufficiently address the audio-visual and perceptual aspects of the concepts they model. As a result, they have limited use for searching, browsing, or summarizing multimedia information repositories.

Please replace the paragraph beginning at page 4, line 8 with the following paragraph:

It is, therefore, an objective of the present invention to provide a method and apparatus for encoding of knowledge using a multimedia network that integrates concepts, relations, words, and multimedia content into a single representation. The multimedia network builds on WORDNET by providing additional signifiers of the semantic concepts using multimedia content and defining perceptual relations based on the features of the multimedia content.

Please replace the paragraph beginning at page 10, line 1 with the following paragraph:

Referring to Figure 2, there is shown an encoding of the media network knowledge representation (111) of the present invention. The media network represents concepts (200, 201, 202) and their signifiers, which may be words (205, 206) and content (207, 208, 209), as nodes. The media network represents relationships, which may be semantic and lexical relationships (210), content relationships (211, 212), and feature relationships as arcs between the nodes. The graphical representation shown in Figure 2 is helpful in visualizing the media network knowledge representation, however, in practice, the media network knowledge representation can be fully represented using any computer data structures that allow modeling of graphs or networks.

Please replace the paragraph beginning at page 11, line 3 with the following paragraph:

Referring to Figure 4, there is shown one example process for creating a media network knowledge representation. This process assumes that a lexical network knowledge representation such as WORDNET is already constructed using a process such as that shown in steps (407) and (408) in which concepts are identified, words are associated with the concepts and the lexical and semantic relationships are encoded. This forms the initial media network knowledge representation (111). The process continues by supplying multimedia content (400) to the classification system in step (401) which classifies the content by associating concepts and words with the content. The classification system can be a manual process in which a human ascribes labels to the content. Alternatively, the classification system can be fully automated in which the content is assigned different labels on the basis of its automatically extracted features. The extraction of features of multimedia is a well-known process in the case of a large number of feature descriptors, such as color histograms, shape signatures, edge direction histograms, or texture descriptors, in which the feature descriptors are generated by processing the multimedia signals. Finally, there are solutions that are semiautomatic in which a human with assistance of a computer classifies the content. Given the results of the classification, in step (403), the content is

attached to the concept nodes of the media network knowledge representation (111).

Please replace the paragraph beginning at page 11, line 19 with the following paragraph:

The process continues by supplying multimedia content (400) to the feature extraction system in step (402) which analyzes the content and extracts descriptors of the audio or visual features of the content. Example features include color, texture, shape, motion, melody, and so forth. Example descriptors include color histogram and Fourier moments. Given the results of the feature extraction, in step (404), the descriptors are associated with the content nodes of the media network knowledge representation. Finally, the descriptors are supplied to the similarity computation system in step (405) which computes the similarity of the content based on the values of the descriptors. The similarity may be derived by computing the distance between the multi-dimensional vectors that represent the feature descriptors. For example, the Euclidean distance metric, walk-metric, or quadratic-form distance metric may be used. The value of the similarity measurement may be used to assign a particular strength to an arc in the multimedia network. This may have the consequence of making some arcs more important than others. Furthermore, multiple arcs may be defined between content nodes in the case that multiple features are described. For example, one arc may correspond to the texture similarity, while another arc may refer to the shape similarity. Furthermore, arcs may also correspond to an integration of features, such as referring to the joint similarity in terms of color and shape. Given the results of the similarity computation, in step (406), the feature similarity is represented as relationships or arcs in the media network knowledge representation (409).

Please replace the paragraph beginning at page 12, line 14 with the following paragraph:

Referring to Figure 5, the media network knowledge representation can be encoded (501) using the ISO MPEG-7 Description Definition Language (DDL) as shown in Table 1 to provide an XML representation

of the media network knowledge representation (111). The MPEG-7 representation can be further encoded into a compact binary form using the MPEG-7 BiM binary encoding system. Once encoded using MPEG-7, the media network knowledge representation (500) can be stored persistently, such as in a database (503), or can be transmitted over a network, or carried with the multimedia data in a transport stream.